SIGNALING PATHWAY FOR BUTANOL PRODUCTION IN *CLOSTRIDIUM BEIJERINCKII* NRRL B-598

Jana Musilová

Master Degree Programme (2), FEEC BUT

E-mail: xmusil57@stud.feec.vutbr.cz

Supervised by: Karel Sedlář

E-mail: sedlar@feec.vutbr.cz

Abstract: In this study, we bring the first dynamic model of signaling pathway for butanol production in sporulating, solvent-producing bacterium, *Clostridium beijerinckii* NRRL B-598. The model allows to study functions of individual genes involved in the production of solvents and acids. We used an open platform Cell Collective for designing the model and stored it under the name 'Signaling Pathway for Butanol Production in Clostridium beijerinckii NRRL B-598' version 1.0 (https://research.cellcollective.org/?dashboard=true#36604:1/signaling-pathway-for-butanol-production-in-clostridium-beijerinckii-nrrl-b598/10) in publicly available repository of Cell Collective.

Keywords: signaling pathway, clostridium, butanol

1 INTRODUCTION

With increasing environmental protection and the dwindling amount of oil, efforts are being made to produce fuels from renewable resources. A promising option seems to be the production of butanol as biofuel by solventogenic bacteria. *Clostridia*, gram-positive, anaerobic and sporulating organisms produce solvents such as aceton, butanol and ethanol (ABE) during their life cycle. The strain *C. beijerinckii* NRRL B-598 is able to produce butanol up to maximum of 7.6 g/l [1] before the start of sporulation. *Clostridia*, mainly model organism *C. acetobutylicum* ATCC 824 are widely studied solvent producers, but different strains are substantially diverse in phenotype, therefore obtained knowledge cannot be used for the strain *C. beijerinckii* NRRL B-598.

2 MATERIALS AND METHODS

2.1 DATA GAIN

The creation of the model is based on data obtained by HPLC and RNA-Seq in the previous study of the strain *C. beijerinckii* NRRL B-598 [1] together with text and database mining for homologies in related organisms. Gene matching and its functional annotation was found using the KEGG encyclopedia and its tool BlastKOALA [2].

2.2 MODEL CREATION

Cell Collective [3] (cellcollective.org) is an open platform, freely available for research community. We chose this software as it allows designing a dynamic model, perform a simulation and an analysis, creating conditions between nodes, sharing results and also final publishing on the website.
As it can be seen in the Figure 1, dynamic model contains 66 components (nodes) and 139 interactions (edges). Orange nodes are labelled as external components, which means the ability to set the activation rate of the element during simulation. Gray nodes are called internal components. Edges are shown as arrows – directed graph which indicate the links between the nodes. Green arrows indicate the activation of the node, red arrows the inactivation. Grey arrows show a certain condition (i.e. sporulation is active only when cell membrane is inactive and sigA is active).

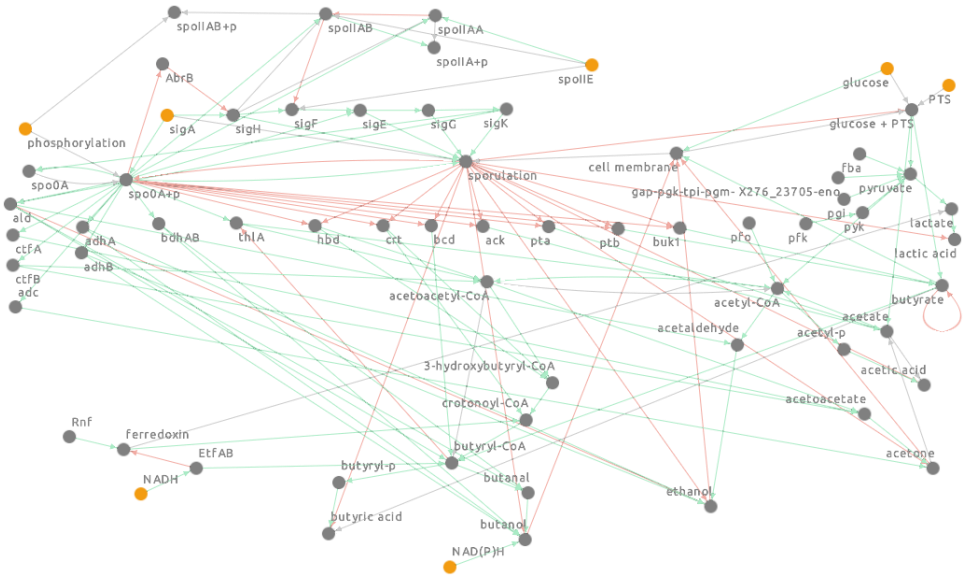

Figure 1: Signaling pathway of butanol production in *C. beijerinckii* NRRL B-598

2.3 SIMULATION

In order to study the process of butanol production and the involvement of individual genes and metabolites, we performed a simulation in Cell Collective. To verify the accuracy of the dynamic model, we compared created chart with real data obtained from HPLC [1] for the following metabolites: acetone, butanol, ethanol, acetic acid, butyric acid, lactic acid and glucose. Whereas real data were measured as a concentration over time, it was necessary to convert them into a production/consumption rate over time. For this purpose, we used the Matlab 2017b and the Equation 1:

$$v_X = \frac{dc_X}{dt} \; [gl^{-1}h^{-1}] \tag{1}$$

where v_X is the reaction rate or activity of a metabolite X, c_X is the concentration a metabolite X, and t is time. The results are shown in the Figure 2.

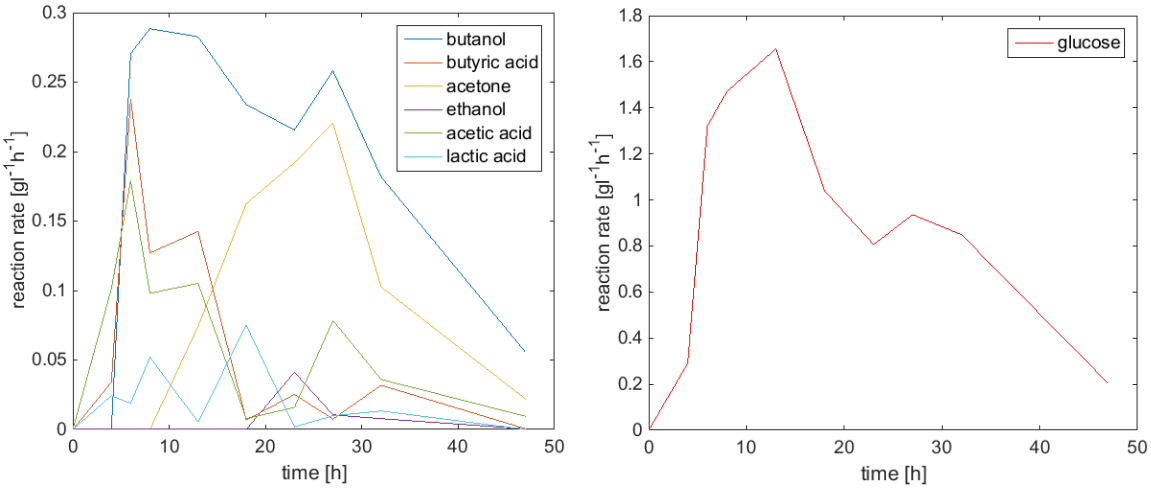

Figure 2: Production/consumption rate of metabolites obtained from HPLC

In the next step, we performed a model fitting by adjusting the activity rate of external components to match the real data shown above. Activity levels of individual external components during the simulation are entered in the Table 1. To get comparative results with HPLC, based on the dynamic analysis, we reduced activity level of glucose from 100 % to 0 % in step 35. Result of the simulation is shown in the Figure 3.

activity [%] time step [-]	glucose	NAD(P)H	NADH	phosphorylation	PTS	sigA	spoIIE
0 - 34	100	50	50	100	100	100	75
35 - 92	0	50	50	100	100	100	75

Table 1: Activity level of external components during simulation

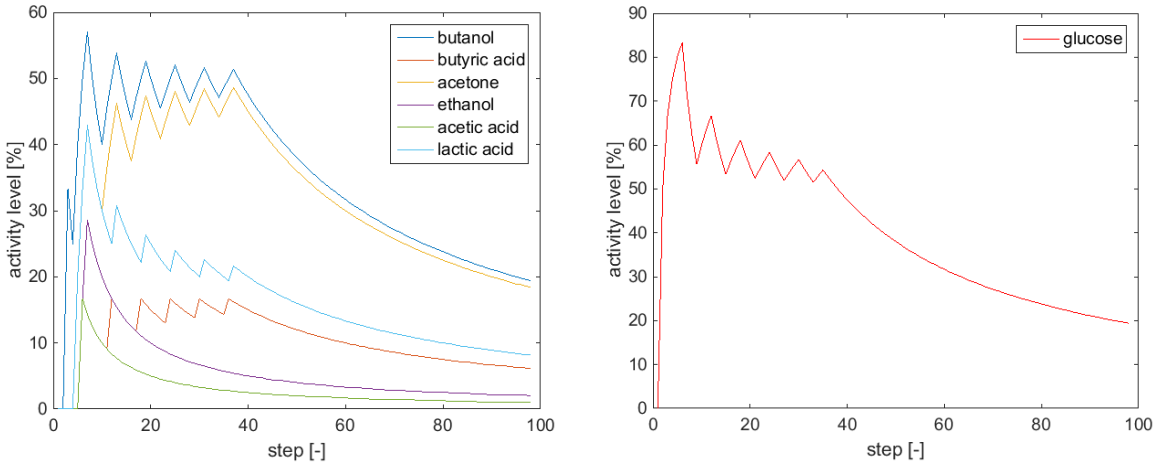

Figure 3: Simulation of the dynamic model in Cell Collective

2.4 RESULTS

Values obtained in the laboratory and converted into a production/consumption rate over time are given in units [$gl^{-1}h^{-1}$] and indicate the rate of production or consumption of a certain metabolite. Values from the simulation are given in percent of activation of a certain metabolite over step, because the model is discrete and therefore the x-axis shift is given in points (steps), not in time. For that reason it is not possible to perform a full statistical evaluation of the results. We compared results of 10 time points and 10 steps using Spearman correlation coefficient and Matlab 2017b function *corr(x, y)*. Values in the individual points and the results of the statistics are shown in the **Table 2**.

time [h], step[-]	butanol		acetone		ethanol	
	production rate [g/l/h]	activity level [%]	production rate [g/l/h]	activity level [%]	production rate [g/l/h]	activity level [%]
0, 1	0	0	0	0	0	0
4, 8	0	50	0	37,5	0	14,3
6, 12	0,27	50	0	41,7	0	8,3
8, 16	0,29	43,8	0	37,5	0	6,3
13, 26	0,28	50	0,07	46,2	0	3,8
18, 36	0,23	50	0,16	47,2	0	2,8
23, 46	0,22	41,3	0,19	39,1	0,04	2,2
27, 54	0,26	35,2	0,22	33,3	0,01	1,9
32, 64	0,18	29,7	0,1	28,1	0,01	1,6
47, 94	0,06	20,2	0,02	19,1	0	1,1
Spearman c. c.	0,6693					

Table 2: Results of simulation and HPLC – values in the individual points and the statistics

However, we can state by comparison of the plots and statistic result that created dynamic model approximates results obtained from the real data. Deviations are mainly caused by the different units in the compared plots. It is also not possible to get results exactly matching real data, as we have created a signaling pathway of butanol production containing 66 components (including genes and metabolites), but the cell *Clostridium beijerinckii* NRRL B-598 contains more than 5 000 genes.
For the exact results, it is necessary to further research of the function of individual genes and to extend the pathway to the genome scale model.

3 CONCLUSION

We created a dynamic model of signaling pathway for butanol production in *Clostridium beijerinckii* NRRL B-598 (see Figure 1) based on data from experiments and already published networks. The pathway is published in the Cell Collective under the name 'Signaling Pathway for Butanol Production in Clostridium beijerinckii NRRL B-598', version 1.0 (https://research.cellcollective.org/?dashboard=true#36604:1/signaling-pathway-for-butanol-production-in-clostridium-beijerinckii-nrrl-b598/10). Based on the model dynamics, there is a possibility (in addition to observe the model) to simulate and analyze the model for studying the pathway as a whole as well as the function of individual components.

To verify the function of the model, we used the real data of the concentration of seven metabolites. We converted the concentration to the production/consumption rate over time and compared the results (see Figure 2) with the dynamic model simulation (see Figure 3) and calculated the Spearman correlation coefficient with the result 0,6693. Real data and model are approximately coincides, but they are not exactly accurate due to the different units in the charts and also due to the model limitations only to the genes and metabolites involved to the butanol production.

ACKNOWLEDGEMENT

This work has been supported by grant project GACR 17-00551S.

REFERENCES

[1] Sedlář, K., et al: Transcription profiling of butanol producer Clostridium beijerinckii NRRL B-598 using RNA-Seq. BMC Genomics, 2018, DOI: 10.1186/s12864-018-4805-8

[2] Kanehisa, M.: BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J. Mol. Biol., 2016, DOI: 10.1016/j.jmb.2015.11.006

[3] Helikar, T., et. al.: The Cell Collective: Toward an open and collaborative approach to systems biology. BMC Syst Biol, 2012. DOI: 10.1186/1752-0509-6-96